

High-resolution bone microstructure imaging based on ultrasonic frequency-domain full-waveform inversion

Yifang Li(李义方), Qinzhen Shi(石勤振), Ying Li(李颖), Xiaojun Song(宋小军), Chengcheng Liu(刘成成), Dean Ta(他得安), and Weiqi Wang(王威琪)

Citation: Chin. Phys. B, 2021, 30 (**1**): 014302. DOI: 10.1088/1674-1056/abc7aa
Journal homepage: http://cpb.iphy.ac.cn; http://iopscience.iop.org/cpb

What follows is a list of articles you may be interested in

Influence exerted by bone-containing target body on thermoacoustic imaging with current injection

Yan-Hong Li(李艳红), Guo-Qiang Liu(刘国强), Jia-Xiang Song(宋佳祥), Hui Xia(夏慧)
Chin. Phys. B, 2019, 28 (**4**): 044302. DOI: 10.1088/1674-1056/28/4/044302

Axial acoustic radiation force on a fluid sphere between two impedance boundaries for Gaussian beam

Yuchen Zang(臧雨宸), Yupei Qiao(乔玉配), Jiehui Liu(刘杰惠), Xiaozhou Liu(刘晓宙)
Chin. Phys. B, 2019, 28 (**3**): 034301. DOI: 10.1088/1674-1056/28/3/034301

Interaction between encapsulated microbubbles: A finite element modelling study

Chen-Liang Cai(蔡晨亮), Jie Yu(于洁), Juan Tu(屠娟), Xia-Sheng Guo(郭霞生), Pin-Tong Huang(黄品同), Dong Zhang(章东)
Chin. Phys. B, 2018, 27 (**8**): 084302. DOI: 10.1088/1674-1056/27/8/084302

Lorentz force electrical impedance tomography using pulse compression technique

Zhi-shen Sun(孙直申), Guo-qiang Liu(刘国强), Hui Xia(夏慧)
Chin. Phys. B, 2017, 26 (**12**): 124302. DOI: 10.1088/1674-1056/26/12/124302

Quantitative damage imaging using Lamb wave diffraction tomography

Hai-Yan Zhang(张海燕), Min Ruan(阮敏), Wen-Fa Zhu(朱文发), Xiao-Dong Chai(柴晓冬)
Chin. Phys. B, 2016, 25 (**12**): 124304. DOI: 10.1088/1674-1056/25/12/124304

High-resolution bone microstructure imaging based on ultrasonic frequency-domain full-waveform inversion*

Yifang Li(李义方)[1,2], Qinzhen Shi(石勤振)[1], Ying Li(李颖)[1], Xiaojun Song(宋小军)[1], Chengcheng Liu(刘成成)[3,†], Dean Ta(他得安)[1,2,3,‡], and Weiqi Wang(王威琪)[1]

[1] *Department of Electronic Engineering, Fudan University, Shanghai, China*
[2] *Human Phenome Institute, Fudan University, Shanghai, China*
[3] *Academy for Engineering and Technology, Fudan University, Shanghai, China*

(Received 21 September 2020; revised manuscript received 2 November 2020; accepted manuscript online 5 November 2020)

The main challenge in bone ultrasound imaging is the large acoustic impedance contrast and sound velocity differences between the bone and surrounding soft tissue. It is difficult for conventional pulse-echo modalities to give accurate ultrasound images for irregular bone boundaries and microstructures using uniform sound velocity assumption rather than getting *a prior* knowledge of sound speed. To overcome these limitations, this paper proposed a frequency-domain full-waveform inversion (FDFWI) algorithm for bone quantitative imaging utilizing ultrasonic computed tomography (USCT). The forward model was calculated in the frequency domain by solving the full-wave equation. The inverse problem was solved iteratively from low to high discrete frequency components via minimizing a cost function between the modeled and measured data. A quasi-Newton method called the limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS) was utilized in the optimization process. Then, bone images were obtained based on the estimation of the velocity and density. The performance of the proposed method was verified by numerical examples, from tubular bone phantom to single distal fibula model, and finally with a distal tibia-fibula pair model. Compared with the high-resolution peripheral quantitative computed tomography (HR-pQCT), the proposed FDFWI can also clearly and accurately presented the wavelength scaled pores and trabeculae in bone images. The results proved that the FDFWI is capable of reconstructing high-resolution ultrasound bone images with sub-millimeter resolution. The parametric bone images may have the potential for the diagnosis of bone disease.

Keywords: quantitative imaging, full-waveform inversion, bone microstructure, ultrasonic computed tomography, high resolution

PACS: 43.60.Lq, 43.80.Qf, 43.35.Wa, 87.63.dh **DOI:** 10.1088/1674-1056/abc7aa

1. Introduction

Bone strength depends on multiple characteristics, such as bone mass, geometrical morphology, and architecture.[1] Bone geometrical morphology and micro-architecture are important features to determine the mechanical properties.[2,3] Osteoporosis is a common bone metabolic disease, typically characterized by the loss of bone mass, reduced thickness, and increased porosity, therefore resulting in an increased risk of fracture.[4] Severe osteoporosis reduces the quality of life, increases the suffering of patients, even endangers patient's life.[4] Consequently, daily monitoring and early diagnosis of bone diseases are of great significance, especially in aging societies.[5] Bone mineral density (BMD) measured by dual energy x-ray absorptiometry (DXA) is the gold standard for evaluating bone status.[6] However, the comprehensive evaluation of bone based on DXA is still challenging, because it is difficult for DXA to quantify bone microstructure and biomechanical properties.[7–9] High-resolution magnetic resonance imaging (HR-MRI) and high-resolution peripheral quantitative computed tomography (HR-pQCT) are two principal modalities of bone imaging with high spatial resolution, which can provide both geometry and microstructure.[10] However, the clinical applications of HR-MRI and HR-pQCT are limited due to the relatively high costs, long examination time, presence of ionizing radiation, and poor portability.[10]

Ultrasound is a mechanical wave, which is physically suited to probe the elastic and structural properties of bone. Quantitative ultrasound (QUS) techniques are deemed to reflect bone mechanical parameters beyond BMD with the advantages of being inexpensive and non-ionizing.[6] Therefore, QUS based techniques were proposed as alternatives to x-ray based methods to characterize bone.[11] QUS approaches are mainly divided into two categories. One is for cortical assessment using ultrasound guided wave (UGW)[12–14] or first arrival signal (FAS)[15] propagation in the long cortical bones. The other is for cancellous measurements, utilizing ultrasonic through-transmission[16] and backscat-

*Project supported by the National Natural Science Foundation of China (Grant Nos. 11827808, 11874289, and 11804056), the National Science Fund for Distinguished Young Scholars of China (Grant No. 11525416), Shanghai Municipal Science and Technology Major Project, China (Grant No. 2017SHZDZX01), Shanghai Talent Development Fund (Grant No. 2018112), State Key Laboratory of ASIC and System Project (Grant No. 2018MS004), and China Postdoctoral Science Foundation (Grant No. 2019M661334).

†Corresponding author. E-mail: chengchengliu@fudan.edu.cn
‡Corresponding author. E-mail: tda@fudan.edu.cn

© 2021 Chinese Physical Society and IOP Publishing Ltd

ter techniques.[17,18] Recently, QUS capitalized on machine learning[19,20] and deep learning[21,22] for bone or porous media assessment. These studies have yielded promising results to predict material parameters, such as thickness,[23,24] porosity,[2,25] bulk velocities,[22,23] phase velocity,[26] etc. However, these studies can hardly provide entire morphological features of bone and ultrasonic imaging of bone tissue remains a challenging task.

The most common method of medical ultrasound imaging makes use of pulse-echo sonography, which generates backscattered reflection images. This modality is suitable for ultrasonic imaging for soft tissues with small heterogeneity, based on the assumption of uniform sound velocity.[27–30] While for the hard tissue (i.e., bone), the speed of sound in bone differs significantly from that of the surrounding soft tissue.[31,32] The pulse-echo method is difficult to accurately image the irregular bone or multilayer structures without *a prior* of sound speed distribution.[33,34] Olofsson[35] applied phase shift migration (PSM) to image regularly layered objects. Some studies further utilized ray-tracing technique[32,33] or modified PSM[36–38] to image the irregularly hard materials. However, the sound velocity distribution must be obtained[33,34] or estimated in advance.[32] It is unrealistic for ultrasonic bone imaging with *a prior* of sound velocity model in the actual scene. Renaud *et al.*[32] achieved *in-vivo* imaging the first section of long bone using prior ranges of sound velocity to iteratively estimate velocities layer by layer. Considering the large acoustic impedance on the cortical bone surface and the porous structure in cancellous part, the strong reflections generated from the boundary and ultrasonic wave attenuation induced by multiple diffraction, scattering, anelastic absorption, and elastic mode-conversions lead to a very complex wavefield for bone imaging.[32,33,39,40] The high amplitude signals from the boundaries may overlap and interfere with the relatively weak signals scattered inside the bone structure. These superimposed effects make it difficult to accurately image inside trabeculae, pores, and the second cross-section.[32] Li *et al.*[41] successfully reconstructed the surface trabeculae of bovine cancellous bone in virtue of ultrasonic backscatter parameters, but cannot present trabecular distribution below surface.

Ultrasonic computed tomography (USCT) is an emerging tool to quantitatively realize parametric imaging using inversion algorithms. The algorithms are computational techniques initially developed to image the interior of the Earth.[42] Tomography techniques are basically divided into three categories. The first one is the travel-time tomography based on ray theory, which uses the first arrival time of ultrasonic waveform to reconstruct the velocity distribution.[43,44] By merely considering refraction, this method is efficient and stable[43] but only valid when the imaging objects are much larger than the wavelength.[45] Its resolution is limited and cannot reconstruct structure below the scale of the diameter in the first Fresnel zone.[39,46] The second method is the diffraction tomography (DT), which takes diffraction and scattering into account with a linearized scattering model such as the Born or Rytov approximations.[47,48] This method is much more time consuming but has higher resolution.[47] The third family is the reconstruction technique called full-waveform inversion (FWI).[45,49,50] FWI attempts to reconstruct the parametric images of the medium by trying to match the measured signals with the results of the full wave equation.[50] FWI is actually the most accurate one in inversion community for allowing higher order diffraction and scattering to be considered in its numerical solver.[51] With a higher cost of computational effort, the FWI provides higher resolution and accuracy than previous two methods that only utilize part of the recorded wavefield.

From travel-time inversion to FWI, the inversion techniques were first used in medical ultrasound imaging for soft tissues.[52–54] In recent years, the inversion methods have been attempted to apply in hard tissue imaging,[39,47] particularly in bone imaging.[11,31,40,55,56] Bone is a heterogeneous, porous, and highly attenuated medium.[55] For using limited information of wavefield, travel-time inversion can only be applied to image macroscopic morphology with relatively low resolution. As the phase transition through the bone tissue is too large, DT based on the first-order Born or Rytov approximation cannot be utilized to tackle the problem with large acoustic impedance misfit.[55] Therefore, Lasaygues *et al.*[40] proposed a compound USCT to address this issue. Some studies[11,47,57] with iterative approaches based on high-order approximations can also image high-contrast targets. A method called distorted Born diffraction tomography (DBDT)[11] was used to image bone-mimicking phantoms, obtaining faithful geometry and sound speed images. With the recorded full wavefield and the experience of imaging the Earth's internal structure, FWI is expected to transfer to hard tissue imaging, further providing higher quality images. Bernard *et al.*[31] gave fairly accurate map of the longitudinal wave speed for cortical bone quantitative imaging in simulation study using time-domain full-waveform inversion (TDFWI). In addition, the approximate density map was also provided to assess the cortical bone. Whereas TDFWI is computationally intensive.

It is understood that the achievable spatial resolution of FWI is approximately half of the wavelength, i.e., $\lambda/2$.[39,58] Using megahertz frequency, the FWI can theoretically image pores and trabeculae in bone tissue with the size of millimeter or sub-millimeter.

As FWI possesses unequivocal advantages in hard tissue imaging via considering the full information of wavefield, the frequency-domain full-waveform inversion (FDFWI) method was introduced into bone parametric imaging in this study. Unlike TDFWI based on the iterative fitting of complete time series between modeled and observed data, the FD-

FWI only depends on the full information of wavefield at some discrete frequencies. This data reduction leads to efficient computation. Besides, FDFWI is more convenient to be directly performed from low to high discrete frequency components, and there is no need for utilizing lowpass filtering, which is a key process for TDFWI to avoid trapping into the local minima.[59,60] To the best of our knowledge, this study is the first report in which the FDFWI method is applied to simultaneously estimate sound velocity and mass density in bone quantitative imaging. The performance of the method is demonstrated by tubular bone phantom, single distal fibula model, and finally with a distal tibia-fibula bone pair model.

The rest of paper is organized as follows. Model assumptions, the theory of FDFWI, and the selected optimization algorithm are stated in Section 2. The numerical simulation setup and the simulated examples are described in Section 3. The results are provided in Section 4. Section 5 gives the discussion, and finally, the conclusions are presented in Section 6.

2. Theory and methods

The objective of this study is to conduct two-dimensional (2-D) parametric imaging of bone cross-sections from ultrasonic signals collected from a ring array transducer. The plane of the ring array probe (i.e., the imaging plane) is orthogonal to the long axis of the bones.

2.1. Problem statement

The tomography parametric imaging is essentially an inverse process to estimate the bone material parameters, such as sound velocity and mass density, which are utilized to quantitatively characterize the bone. The parametric models (i.e., sound velocity and mass density) are discretized as a $M \times N$ grid.

The ultrasound tomography acquisition setup is depicted in Fig. 1. The bone is immersed in water and surrounded by a ring array transducer with n elements. The positions of the elements in the ring array transducer are determined by the location vectors $\boldsymbol{r}_i$ $(1 \leqslant i \leqslant n)$. Each element in the ring array transmits ultrasonic pulses in sequence, and all elements receive signals at the same time. For FDFWI, the transmitting pulse $S(\boldsymbol{r}_{s_i}, \omega)$ $(1 \leqslant s_i \leqslant n)$ is the excitation source in the frequency domain, and the received complex wavefield at the receiver position $\boldsymbol{r}_{r_i}$ $(1 \leqslant r_i \leqslant n)$ is described by $D_{\text{obs}}(\boldsymbol{r}_{r_i}, \omega)$. The measurements acquired from all emitter-receiver pairs can be stacked into a $n \times n$ dimensional matrix $\boldsymbol{D}_{\text{obs}}(\omega)$. Similar to the measured wavefield, the simulated wavefield denoted as $U_{\text{sim}}(\boldsymbol{r}_{r_i}, C, \rho, \omega)$ obtained by numerical calculation at the receiver location $\boldsymbol{r}_{r_i}$ $(1 \leqslant r_i \leqslant n)$ with bone material parameter models (i.e., $C(\boldsymbol{r}), \rho(\boldsymbol{r})$) is also stacked into a $n \times n$ dimensional matrix $\boldsymbol{U}_{\text{sim}}(C, \rho, \omega)$. FDFWI is typically formulated as a least square minimization problem, and the cost function satisfies the following relationships:[61]

$$E(C, \rho, \omega) = \frac{1}{2} \left\| \boldsymbol{U}_{\text{sim}}(C, \rho, \omega) - \boldsymbol{D}_{\text{obs}}(\omega) \right\|^2, \tag{1a}$$

$$\{C, \rho\} = \arg\min_{C,\rho} \frac{1}{2} \left\| \boldsymbol{U}_{\text{sim}}(C, \rho, \omega) - \boldsymbol{D}_{\text{obs}}(\omega) \right\|^2. \tag{1b}$$

The process of FDFWI imaging is to update the model parameters iteratively to approximate the true model values. When the optimization problem converges well, the simulated wavefield $\boldsymbol{U}_{\text{sim}}(C, \rho, \omega)$ best matches the measured wavefield $\boldsymbol{D}_{\text{obs}}(\omega)$. With vectorization of the matrix, the cost function (1) can be rewritten as[59]

$$E(C, \rho, \omega) = \frac{1}{2} \boldsymbol{e}^{\text{H}}(C, \rho, \omega) \boldsymbol{e}(C, \rho, \omega), \tag{2a}$$

$$\boldsymbol{e}(C, \rho, \omega) = \boldsymbol{U}_{\text{sim}}(C, \rho, \omega) - \boldsymbol{D}_{\text{obs}}(\omega), \tag{2b}$$

where H denotes the Hermitian transpose, and $\boldsymbol{e}(C, \rho, \omega)$ is the residual vector between the modeled and measured wavefields. It is worth noting that the wavefield is explicitly depicted in terms of frequency ω and parametric models (i.e., $C(\boldsymbol{r}), \rho(\boldsymbol{r})$). The optimization is performed on one discrete frequency component at a time, and the numerical simulated wavefield depends on the parametric models. When the parametric models are iteratively updated in FDFWI, the numerical simulated wavefield is also updated with the new parametric models.

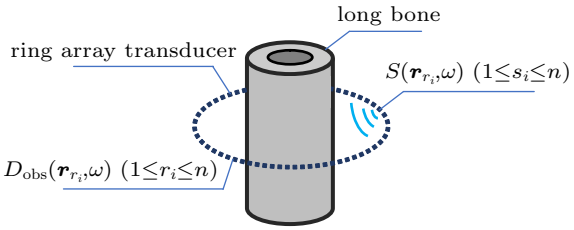


Fig. 1. Ultrasound tomography transducer acquisition setup.

2.2. Forward problem

The propagation of ultrasonic wave is controlled by the acoustic wave equation. Forward modeling of ultrasonic propagation in frequency domain can be described by the Helmholtz equation with variant density[62]

$$-\frac{\omega^2}{K(\boldsymbol{r})} P(\boldsymbol{r}, \omega) - \nabla \cdot \left(\frac{1}{\rho(\boldsymbol{r})} \nabla P(\boldsymbol{r}, \omega) \right) = S(\boldsymbol{r}, \omega), \tag{3a}$$

$$K(\boldsymbol{r}) = \rho(\boldsymbol{r}) C(\boldsymbol{r})^2, \tag{3b}$$

where $\boldsymbol{r}$ is the spatial position vector, $P(\boldsymbol{r}, \omega)$ is the pressure wavefield, ∇ is the gradient operator, $S(\boldsymbol{r}, \omega)$ is a point source, $K(\boldsymbol{r})$ is the spatially-varying bulk modulus, $\rho(\boldsymbol{r})$ is the mass density, and $C(\boldsymbol{r})$ is the compressional wave velocity.

To perform forward numerical calculation, the model is discretized on a $M \times N$ grid. To balance the accuracy and efficiency of the numerical calculation, the grid size is set to be slightly less than $\lambda_{\min}/5$.[61] $\lambda_{\min}$ is the smallest wavelength and is computed using the sound velocity in water at the maximum frequency component. Similar to the parametric model, the forward wavefield $P(\boldsymbol{r}, \omega)$ is sampled as a

$M \times N$ grid at each frequency ω. The discrete pressure values are reorganized as a $M \times N$ dimensional vector P. Similarly, the source $S(\boldsymbol{r}, \omega)$ is also vectorized as a $M \times N$ dimensional vector S. The vector S is sparse and has only one non-zero value whose index corresponds to the position of the emitter at the grid. With the wavefield and the wave source vectorized, the Helmholtz operator in Eq. (3a) is discretized, assembling a sparse and large matrix $\boldsymbol{H}$ with dimension of $MN \times MN$. Matrix $\boldsymbol{H}$ is actually the numerical approximation of the Helmholtz operator underlying partial differential stencil. The numerical construction of $\boldsymbol{H}$ and the absorbing boundary can refer to the work detailed in Ref. [62]. After discretization, the Helmholtz Eq. (3) can be rewritten in matrix form

$$\boldsymbol{H}(C, \rho, \omega) P(\omega) = S(\boldsymbol{r}_{s_i}, \omega), \quad (1 \leqslant s_i \leqslant n), \tag{4}$$

where $\boldsymbol{H}$ may be a high-dimensional matrix according to the specific requirement of model discretization. Direct inverse computation of high-dimensional matrix is time-consuming and inefficient. To solve Eq. (4) for a multiple source problem, matrix factorization methods such as LU decomposition are strongly recommended for increasing the computation efficiency.[63] When LU decomposition is used to solve Eq. (4) once for one source, LU constituents of $\boldsymbol{H}$ can be stored to rapidly solve wavefield P for other sources. This is critical to save the computational costs of iterative solutions for inverse problems with multiple sources.

2.3. Inverse problem and the adjoint-state method

Since there are large data quantities and model parameters involved in FDFWI, local optimization methods[42,50] based on gradient descent are generally used to solve the cost function (2). Accurately calculating the gradient of the cost function in iterative process is a key point for the success of inversion[42,61]

$$\nabla_C E(C, \rho, \omega) = \boldsymbol{J}_C^{\mathrm{H}}(C, \rho, \omega) \boldsymbol{e}(C, \rho, \omega), \tag{5a}$$

$$\nabla_\rho E(C, \rho, \omega) = \boldsymbol{J}_\rho^{\mathrm{H}}(C, \rho, \omega) \boldsymbol{e}(C, \rho, \omega), \tag{5b}$$

$$\boldsymbol{J}_C(C, \rho, \omega) = \frac{\delta \boldsymbol{e}(C, \rho, \omega)}{\delta C} = \frac{\delta \boldsymbol{U}_{\mathrm{sim}}(C, \rho, \omega)}{\delta C}, \tag{5c}$$

$$\boldsymbol{J}_\rho(C, \rho, \omega) = \frac{\delta \boldsymbol{e}(C, \rho, \omega)}{\delta \rho} = \frac{\delta \boldsymbol{U}_{\mathrm{sim}}(C, \rho, \omega)}{\delta \rho}, \tag{5d}$$

where the gradients ∇_C and ∇_ρ are taken with respect to the sound velocity and mass density, respectively. The gradient formulations in Eq. (5) are matrix-vector products. J_C and J_ρ are the Jacobian matrices of the forward problem that contains the Fréchet derivatives of the received wavefield with respect to model parameters (i.e., C, ρ). $e(C, \rho, \omega)$ is the residual vector that is the same as that in Eq. (2).

The inversion is conducted step by step from low to high frequencies from a selected set of frequency components. Given the current estimations of the parameter models $C^{(i)}(\omega), \rho^{(i)}(\omega)$ and the current gradients $\nabla_C E(C^{(i)}, \rho^{(i)}, \omega), \nabla_\rho E(C^{(i)}, \rho^{(i)}, \omega)$, the updated parameter models with the normal gradient descent method can be written as

$$C^{(i+1)}(\omega) = C^{(i)}(\omega) - \alpha \nabla_C E(C^{(i)}, \rho^{(i)}, \omega), \tag{6a}$$

$$\rho^{(i+1)}(\omega) = \rho^{(i)}(\omega) - \alpha \nabla_\rho E(C^{(i)}, \rho^{(i)}, \omega), \tag{6b}$$

where α is the step size which could be a constant or determined by a line search method.[64] Nevertheless, the gradient computation in such an FWI problem including multiple parameters and high-dimensional matrix is computationally challenging. If the gradients in Eq. (5) are directly calculated, the computations are so intensive. When the model parameter in one grid point is disturbed for differential computation every time, a forward calculation is needed to obtain the recorded wavefield. The forward calculation requires performing as many as the number of model parameters for each source.[50,65] Therefore, a total of $n \times M \times N$ times of forward calculation is needed to get $\boldsymbol{J}_C$ and $\boldsymbol{J}_\rho$ for n sources per iteration. When the model parameters are large, the calculation of this part takes a huge amount of time. The dimension of $\boldsymbol{J}_C$ and $\boldsymbol{J}_\rho$ is $n^2 \times MN$. Obviously, it is too time-consuming for bone quantitative imaging in practical use.

Fortunately, there is an alternative way called the adjoint state method[65] that does not compute the Jacobian matrix explicitly to get the gradient. This method was developed to efficiently compute the gradient. In the numerical community, it is now a well-known method for gradient calculation of a functional with respect to the model parameters. The functional depends on model parameters through state variables, which are solutions of the forward problem.[42] The adjoint state corresponds to the wavefield emitted and back-propagated from the receivers, which is deemed independent of the model parameters.[66] This approach is often implemented resorting to an augmented functional also called associated Lagrangian. It is actually an optimization problem with partial differential equation (PDE) constrained in this study. The augmented functional is defined as

$$\begin{aligned} L(P, V_{\mathrm{a}}, C, \rho, \omega) &= \mathrm{Re}\left[\frac{1}{2}\left\|U_{\mathrm{sim}}(C, \rho, \omega) - D_{\mathrm{obs}}(\omega)\right\|^2 - \sum_{s_i=1}^{n} \left\langle V_{\mathrm{a}}(\boldsymbol{r}_{s_i}, \omega), \boldsymbol{H}(C, \rho, \omega) P(\boldsymbol{r}_{s_i}, \omega) - S(\boldsymbol{r}_{s_i}, \omega)\right\rangle\right] \\ &= \mathrm{Re}\left[\frac{1}{2}\sum_{s_i=1}^{n}\left\|U_{\mathrm{sim}}(\boldsymbol{r}_{s_i}, C, \rho, \omega) - D_{\mathrm{obs}}(\boldsymbol{r}_{s_i}, \omega)\right\|^2 - \sum_{s_i=1}^{n} \left\langle V_{\mathrm{a}}(\boldsymbol{r}_{s_i}, \omega), \boldsymbol{H}(C, \rho, \omega) P(\boldsymbol{r}_{s_i}, \omega) - S(\boldsymbol{r}_{s_i}, \omega)\right\rangle\right] \end{aligned}$$

$$= \mathrm{Re}\left[\frac{1}{2}\sum_{s_i=1}^{n}\|\boldsymbol{C}\boldsymbol{h}_{\mathrm{r}}P(\boldsymbol{r}_{s_i},\omega)-D_{\mathrm{obs}}(\boldsymbol{r}_{s_i},\omega)\|^2-\sum_{s_i=1}^{n}\langle V_{\mathrm{a}}(\boldsymbol{r}_{s_i},\omega),\boldsymbol{H}(C,\rho,\omega)P(\boldsymbol{r}_{s_i},\omega)-S(\boldsymbol{r}_{s_i},\omega)\rangle\right]$$

$$= \mathrm{Re}\left[\frac{1}{2}\sum_{s_i=1}^{n}\|\boldsymbol{C}\boldsymbol{h}_{\mathrm{r}}P(\boldsymbol{r}_{s_i},\omega)-D_{\mathrm{obs}}(\boldsymbol{r}_{s_i},\omega)\|^2-\sum_{s_i=1}^{n}V_{\mathrm{a}}(\boldsymbol{r}_{s_i},\omega)^{\mathrm{H}}(\boldsymbol{H}(C,\rho,\omega)P(\boldsymbol{r}_{s_i},\omega)-S(\boldsymbol{r}_{s_i},\omega))\right], \quad (7)$$

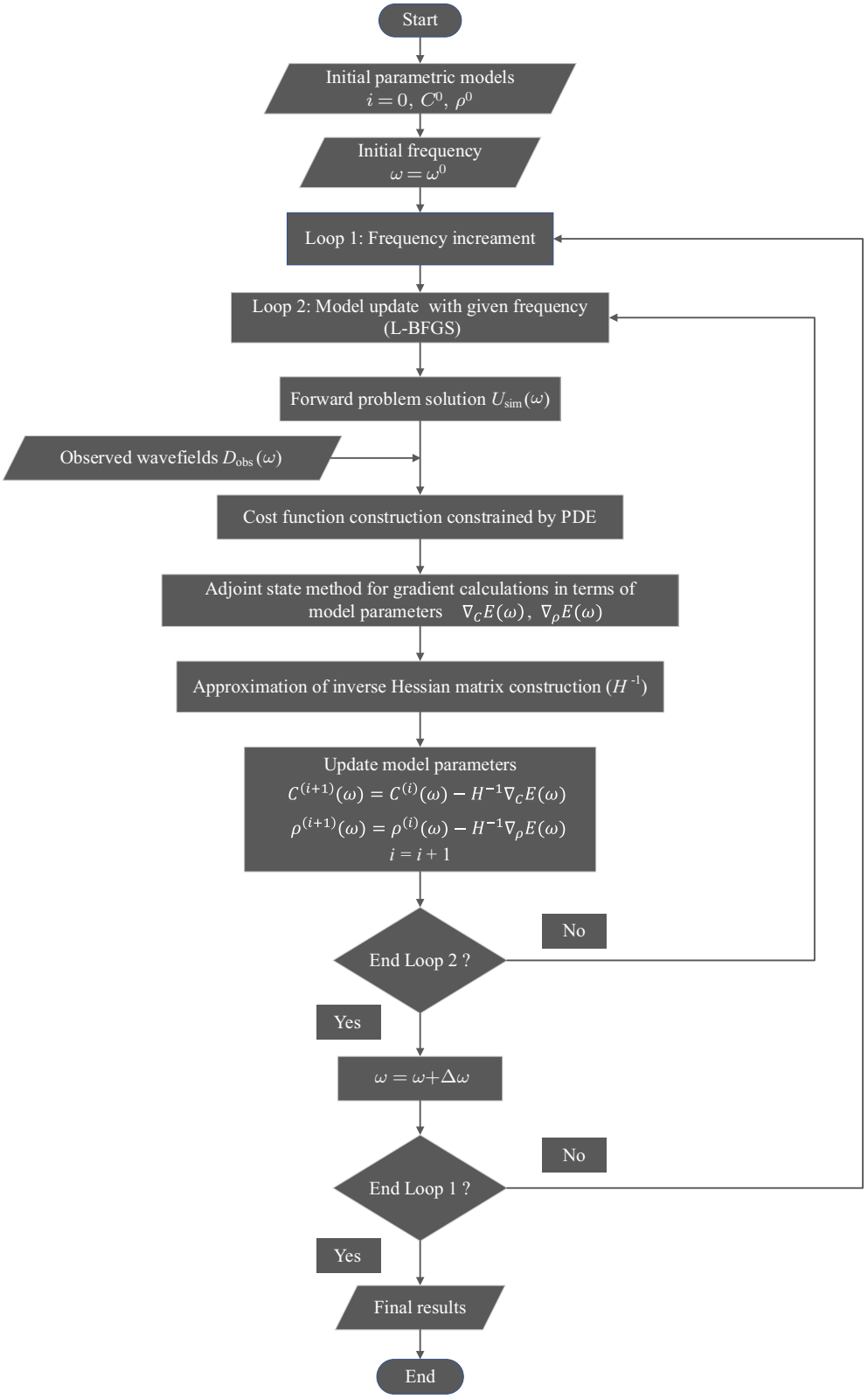


Fig. 2. Flow chart of the FDFWI.

where $\langle\cdot\rangle$ denotes the scalar product of $M\times N$ dimensional complex vectors and $\boldsymbol{r}_{s_i}$ $(1\leqslant s_i\leqslant n)$ depicts the position of the source. $U_{\rm sim}(\boldsymbol{r}_{s_i},C,\rho,\omega)$ and $D_{\rm obs}(\boldsymbol{r}_{s_i},\omega)$ show the simulated and measured wavefields generated by the source $S(\boldsymbol{r}_{s_i},\omega)$ that situated at $\boldsymbol{r}_{s_i}$. Note that $V_{\rm a}(\boldsymbol{r}_{s_i},\omega)$ and $P(\boldsymbol{r}_{s_i},\omega)$ represent the adjoint and forward wavefield vectors related to the source $S(\boldsymbol{r}_{s_i},\omega)$ located at $\boldsymbol{r}_{s_i}$. They are both $M\times N$ dimensional vectors. $\boldsymbol{Ch}_{\rm r}$ is the selective matrix onto positions of the receivers for all sources, with the dimension of $n\times MN$. $V_{\rm a}(\boldsymbol{r}_{s_i},\omega)$ is the adjoint state and also the Lagrange multiplier per source and per frequency component, which is defined by

$$\begin{aligned}\frac{\delta L(\boldsymbol{r}_{s_i},P,V_{\rm a},C,\rho,\omega)}{\delta P(\boldsymbol{r}_{s_i},\omega)} &= \boldsymbol{Ch}_{\rm r}^{\rm H}(\boldsymbol{Ch}_{\rm r}P(\boldsymbol{r}_{s_i},\omega)-D_{\rm obs}(\boldsymbol{r}_{s_i},\omega))\\ &\quad-\boldsymbol{H}^{\rm H}(C,\rho,\omega)V_{\rm a}(\boldsymbol{r}_{s_i},\omega)=0,\end{aligned}\tag{8}$$

then we can obtain

$$\begin{aligned}V_{\rm a}(\boldsymbol{r}_{s_i},\omega) &= \boldsymbol{H}^{-{\rm H}}(C,\rho,\omega)\boldsymbol{Ch}_{\rm r}^{\rm H}(\boldsymbol{Ch}_{\rm r}P(\boldsymbol{r}_{s_i},\omega)\\ &\quad-D_{\rm obs}(\boldsymbol{r}_{s_i},\omega)).\end{aligned}\tag{9}$$

The gradients of L with respect to each model parameter (i.e., C_i,ρ_i $(1\leqslant i\leqslant MN)$) can be calculated by

$$\begin{aligned}&\frac{\delta L(P,V_{\rm a},C,\rho,\omega)}{\delta C_i}\\ &= -{\rm Re}\sum_{s_i=1}^{n}(V_{\rm a}^{\rm H}(\boldsymbol{r}_{s_i},\omega)\frac{\delta\boldsymbol{H}(C,\rho,\omega)}{\delta C_i}P(\boldsymbol{r}_{s_i},\omega)),\end{aligned}\tag{10a}$$

$$\begin{aligned}&\frac{\delta L(P,V_{\rm a},C,\rho,\omega)}{\delta \rho_i}\\ &= -{\rm Re}\sum_{s_i=1}^{n}(V_{\rm a}^{\rm H}(\boldsymbol{r}_{s_i},\omega)\frac{\delta\boldsymbol{H}(C,\rho,\omega)}{\delta \rho_i}P(\boldsymbol{r}_{s_i},\omega)).\end{aligned}\tag{10b}$$

If equation (9) is substituted into Eq. (10), the final gradients will be obtained as

$$\begin{aligned}\frac{\delta L}{\delta C_i} &= -{\rm Re}\sum_{s_i=1}^{n}(P^{\rm H}(\boldsymbol{r}_{s_i},\omega)\frac{\delta^{\rm H}\boldsymbol{H}(\omega)}{\delta C_i}\boldsymbol{H}^{-{\rm H}}(\omega)\\ &\quad\times\boldsymbol{Ch}_{\rm r}^{\rm H}(\boldsymbol{Ch}_{\rm r}P(\boldsymbol{r}_{s_i},\omega)-D_{\rm obs}(\boldsymbol{r}_{s_i},\omega))),\end{aligned}\tag{11a}$$

$$\begin{aligned}\frac{\delta L}{\delta \rho_i} &= -{\rm Re}\sum_{s_i=1}^{n}(P^{\rm H}(\boldsymbol{r}_{s_i},\omega)\frac{\delta^{\rm H}\boldsymbol{H}(\omega)}{\delta \rho_i}\boldsymbol{H}^{-{\rm H}}(\omega)\\ &\quad\times\boldsymbol{Ch}_{\rm r}^{\rm H}(\boldsymbol{Ch}_{\rm r}P(\boldsymbol{r}_{s_i},\omega)-D_{\rm obs}(\boldsymbol{r}_{s_i},\omega))),\end{aligned}\tag{11b}$$

where $\sum_{s_i=1}^{n}[\cdot]$ means that the final gradient for each iteration needs to sum the results of all sources for single model parameter. The computation of $\delta L/\delta C_i$, $\delta L/\delta\rho_i$ with Eqs. (9), (10), and (11) is called the adjoint-state method.[42,66] With this method, the number of forward computations to obtain the gradients is reduced to two for each source per iteration: one to compute the incident wavefield P and the other to back propagate the corresponding residuals to get $V_{\rm a}$. For the convenience of calculation, the gradients corresponding to different parameters are combined into an $M\times N$ dimensional gradient vector to update the model parameters. Figure 2 outlines the flow chart of the FDFWI algorithm.

2.4. Optimization algorithm

There are many local optimization methods based on gradient descent, for instance, the steepest descent, conjugate gradient, Newton, Gauss–Newton, and quasi-Newton.[67] The convergence speed of the steepest descent is slow, which is not suitable for the large data optimization problem in FWI. The convergence rate of Newton and Gauss-Newton is faster than the conjugate gradient and the steepest descent, leading to more stable inversion process.[67] However, the computation and storage cost of the Hessian or approximate Hessian matrix is so intensive in iterative process (i.e., $X_{k+1}=X_k-\boldsymbol{H}^{-1}\nabla_{X_k}E$, where X_k denotes the model parameter, and $\boldsymbol{H}$ denotes the Hessian matrix). $\boldsymbol{H}^{-1}$ indicates the search direction. It is usually necessary to avoid the direct inverse of such large matrix, especially for the strong nonlinear and large-scale problem in FWI. Therefore, a matrix-free variant of the quasi-Newton method known as limited-memory Broyden–Fletcher–Goldfarb–Shanno algorithm (L-BFGS)[68] is utilized in this paper, allowing constructing an approximation of the inverse Hessian in a recursive way without explicitly forming it. Only a few gradients of the previous nonlinear iterations, typically 3–20 iterations, need to be stored in L-BFGS to compute the approximate Hessian. Here we choose $m=20$. The L-BFGS method represents a negligible storage and computational cost compared to the conjugate-gradient algorithm,[50] and a more efficient convergence rate than the steepest descent.[31] The implementation of L-BFGS algorithm is detailed in the literature.[68]

2.5. Regularization

FDFWI for bone imaging is a nonlinear and ill-posed inverse problem. The most common way to alleviate the inverse instability and ill-posedness caused by the high-contrast structures[50,69] and the presence of noises[31] is to use regularization. Tikhonov and total variation (TV) regularization are well known in the inversion process. The former is helpful to restore the continuity and smoothness of the images.[70] Whereas the latter is able to preserve the sharpness of the edges.[71] The bone parametric imaging requires not only to maintain the continuity and smoothness of some areas, such as soft tissue and bone, but also to restore some regions with high sound velocity contrast, for instance sharp boundaries between bone and soft tissue. Therefore, we tried to combine the advantages of the above two regularizations. In this paper, the Sobolev space norm was employed as a regularization term to keep a balance between sharpness and smoothness in imaging. Mathematically, it is depicted by[72]

$$J_{\rm SN}^{C_i}=(\nabla C_i\cdot\nabla C_i+\varepsilon)^{p/2},\quad\varepsilon>0,\tag{12a}$$

$$J_{\rm SN}^{\rho_i}=(\nabla\rho_i\cdot\nabla\rho_i+\varepsilon)^{p/2},\quad\varepsilon>0.\tag{12b}$$

The Sobolev space norm can easily be added to the cost function with weight λ_{SN}, and the gradient of regularization term (13) is added to the gradient of the cost function (11) for model parameter update. The mathematical features of Sobolev space norm can refer to the literature.[72] In brief, it is a natural blend of Tikhonov and TV regularization.[72]

$$\frac{\delta J_{\mathrm{SN}}^{C_i}}{\delta C_i} = p\nabla \cdot \frac{\nabla C_i}{(\nabla C_i \cdot \nabla C_i + \varepsilon)^{1-p/2}}, \quad \varepsilon > 0, \tag{13a}$$

$$\frac{\delta J_{\mathrm{SN}}^{\rho_i}}{\delta \rho_i} = p\nabla \cdot \frac{\nabla \rho_i}{(\nabla \rho_i \cdot \nabla \rho_i + \varepsilon)^{1-p/2}}, \quad \varepsilon > 0, \tag{13b}$$

where ε is utilized to avoid being divided by zero when calculating the gradient of the regularization term. As long as the value is small enough, it will have little impact on the restored images. We selected $\varepsilon = 10^{-7}$, $\lambda_{\mathrm{SN}} = 10^{-4}$, and $P = 1.5$ in this work. The sum of Eqs. (11) and (13) gives the gradient of a single model parameter with the regularization term. In the iterative calculation, the gradients were combined into an $M \times N$ dimensional vector to update the model parameters.

3. Numerical simulations setup

FDFWI is a nonlinear and ill-posed inverse problem, meaning that an infinite number of models can match the data, particularly when starting inversion with high frequency.[50] Therefore, the initial frequency must be low enough. If not, the so-called cycle-skipping effect will result in the convergence to a local minima.[50] To avoid this effect, the TDFWI proposes consecutive inversion with gradually increasing frequency component using lowpass filtering, as low frequencies are insensitive to cycle-skipping effects.[31] However, the FDFWI provides a more natural workflow for this multiscale approach by performing successive inversions from low to high discrete frequency components.[59,60]

For simplicity, the transmitting sources are ideal point sources in this study. The source waveform is wide-band Ricker wavelet with a dominant frequency f_{c}

$$S(t) = \left[1 - 2\pi^2 f_{\mathrm{c}}^2 (t - t_0)^2\right] \mathrm{e}^{-\pi^2 f_{\mathrm{c}}^2 (t-t_0)^2}. \tag{14}$$

The time domain waveforms and amplitude spectrum of the Ricker wavelet pulse are shown in Fig. 3. Two sources with different center frequencies were used to cover the bandwidth we need. For the transmitting source with $f_{\mathrm{c}} = 200$ kHz, the utilized bandwidth is from 100 kHz to 400 kHz. For the transmitting source with $f_{\mathrm{c}} = 1.75$ MHz, the used bandwidth ranges from 400 kHz to 3.5 MHz. The sound speed and mass density in bone tissue were derived from previous literature,[73] in which the sound speed is 2800 m/s and the mass density is 1800 kg/m^3. For the surrounding media (i.e. ,water or soft tissues), the sound velocity and mass density are set to 1500 m/s and 1000 kg/m^3. The spatial size (i.e., grid size) in the numerical model was calculated by the maximum frequency component and the background sound velocity 1500 m/s. 128 sources and 128 receivers were utilized to produce and receive the ultrasonic signals.

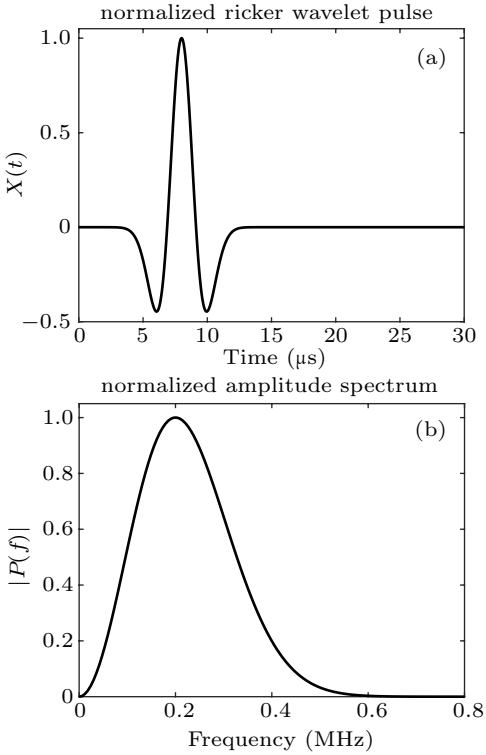


Fig. 3. Ricker wavelet source. (a) Time-domain signal. (b) Amplitude spectrum, with a central frequency $f_{\mathrm{c}} = 200$ kHz.

3.1. Tubular bone phantom

In the first numerical model, a simple tubular bone phantom with a thickness of 2 mm was used, situating in the center of the ring array. The grid size Δh was 60 μm and the 18 mm×18 mm model was therefore discretized as 301×301 grids. The diameter of the ring array transducer was 17 mm. In this case, the mass density model was known and only the sound velocity model was estimated. The speed of sound in water (1500 m/s) was utilized as the initial guess of the velocity model. The simulated data was noiseless, so regularization based on Sobolev space norm was closed in this section, as well as in Subsections 3.2 and 3.3. The frequency of inversion was initially set to 100 kHz and then gradually increased to 3.5 MHz with a step of 100 kHz.

3.2. Single distal fibula model

In the second numerical model, a single distal fibula model was utilized. The model was made by binarization of a HR-pQCT image of distal fibula (Figs. 5(a) and 6(a)). The sizes of the discretized model and the circular ring transducer were the same as those in tubular bone phantom. However, in this circumstance, the sound velocity and mass density were

simultaneously estimated. The frequency components were set from 100 kHz to 3.5 MHz with an interval of 50 kHz. For low frequencies are essential for the reconstruction of mass density,[50,74] the smaller interval can ensure enough components with relatively low frequency in inversion. Both the initial sound velocity and mass density model were uniform, in which the velocity was set to 1500 m/s and the mass density was 1000 kg/m^3 (physical parameters in water).

3.3. Distal tibia-fibula pair model

In the third numerical model, we used a distal tibia-fibula pair model. The model was also derived from binarization of a HR-pQCT image (Figs. 9(a) and 10(a)). As presented in Subsections 4.1 and 4.2, since the inversion results of the first two models verified that the FDFWI can achieve a good reconstruction when the frequency reached 2.5 MHz, the frequency set here was selected from 100 kHz to 2.5 MHz with an interval of 50 kHz. The grid size was set to 100 μm and the 60 mm×60 mm model was discretized as 601 × 601 grids. The diameter of the ring array transducer was 58 mm. This was a challenging scene involving multiple scattering between the two bones. The complex distribution of trabeculae in the distal tibia would introduce significant multiple diffraction, scattering, and reverberation effects. The initial values of velocity and density model were identical with those in the single distal fibula model.

3.4. Robustness against noise

To investigate the robustness of FDFWI against noise, Gaussian random noise was added to the synthetic data generated in the distal tibia-fibula pair model for velocity and density inversion. The reference pressure was defined as the mean amplitude of the received complex signals. The cases with signal-to-noise ratio (SNR) of 30 dB, 10 dB, and 0 dB were respectively performed and discussed. The settings of inversion were consistent with those in the previous Subsection 3.3, except that the Sobolev space norm was utilized. It should be noted that the regularized weight λ_{SN} represents the ratio between the data fidelity item and the penalty term. $\lambda_{SN} = 10^{-4}$ was chosen via the qualitative evaluation of the final inversion results.

4. Results

4.1. Tubular bone phantom

The inversion results with different maximum frequencies are shown in Fig. 4. When the inversion frequency gradually increases to 1.5 MHz, both the outer and inner edges of the phantom are recovered. When the frequency reaches 2.5 MHz, the image becomes clearer with fewer artifacts. However, as the frequency ranging from 2.5 MHz to 3.5 MHz, the improvement of image quality is relatively limited. The result shows that the FDFWI has ability to accurately restore the macroscopic morphology. The inversion error of the velocity is small. If the imaging region in the ring array transducer is selected as the region-of-interest (ROI), the root-mean-square-error (RMSE) of sound velocity in ROI is 149.76 m/s, and the mean relative error in ROI is 6.97%. The RMSE exclusively in bone area is 250.52 m/s, and the mean relative error is 8.95%.

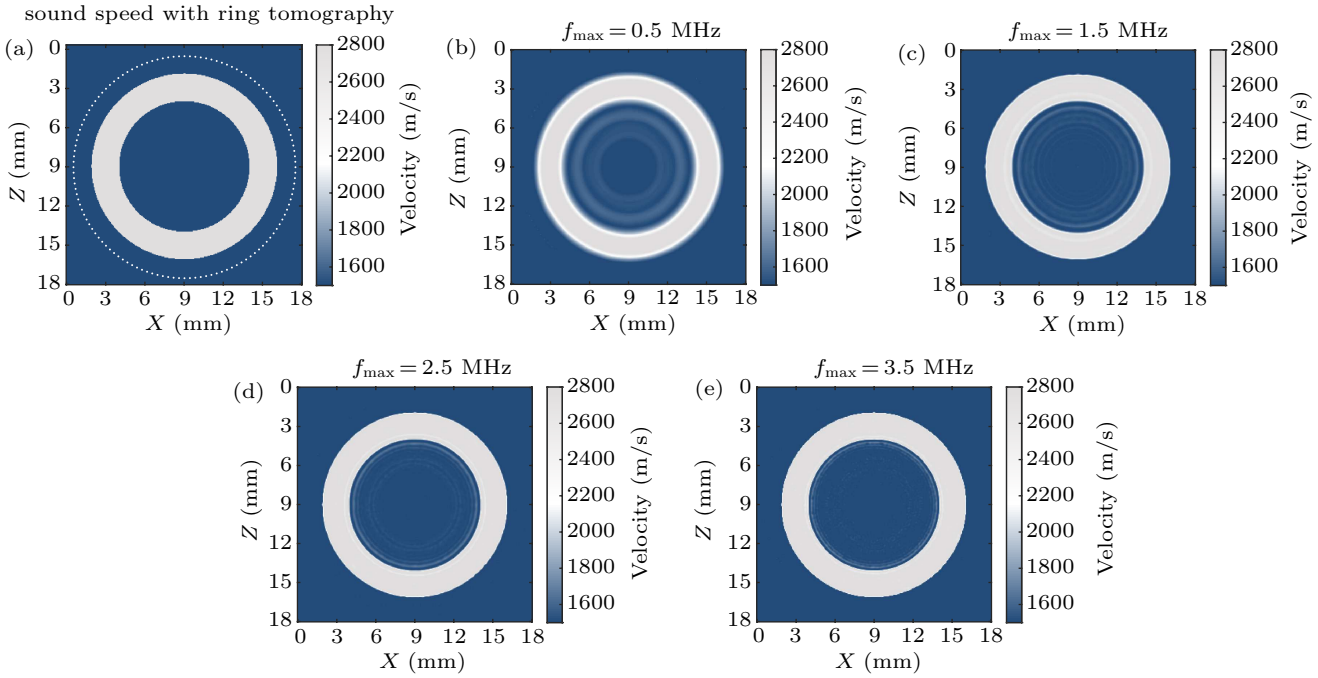


Fig. 4. FDFWI results of the tubular bone phantom for velocity model with different maximum frequencies. (a) True velocity model with ring array transducer. (b)–(e) Result obtained with final frequency of 0.5 MHz, 1.5 MHz, 2.5 MHz, and 3.5 MHz, respectively.

4.2. Single distal fibula model

The inversion results with different maximum frequencies are shown in Figs. 5 and 6. When the inversion frequency gradually increases to 1.5 MHz, both the outer and inner edges of the velocity map are recovered accurately. Meanwhile, partial microstructures in the fibula, such as pores and trabeculae, can be clearly depicted. When the frequency reaches 2.5 MHz, the velocity map becomes clearer and some subtler features can be precisely displayed. Similar to the result of tubular bone phantom, as for the frequency ranging from 2.5 MHz to 3.5 MHz, the improvement of the image quality is relatively limited. The result shows that the FDFWI can provide high-resolution images for fibula bone. Both the geometry and microstructure can be clearly and accurately recovered. Due to the simultaneous inversion of two parameters, the inversion task is relatively more complex. The error of sound velocity increases compared with the single speed inversion in the previous section. For the sound velocity, RMSE in ROI is 258.31 m/s, and the mean relative error in ROI is 12.01%. RMSE in the bone area is 408.39 m/s, and the mean

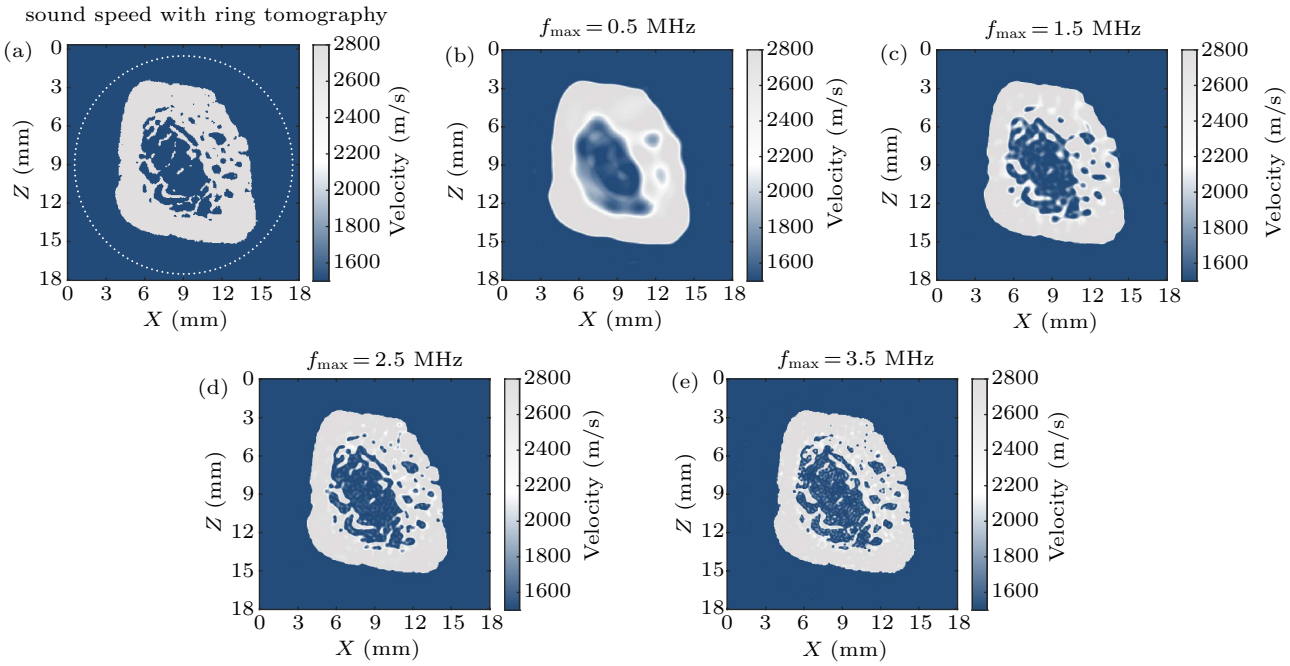


Fig. 5. FDFWI results of the single distal fibula for velocity model with different maximum frequencies. (a) True velocity model (derived from HR-pQCT image). (b)–(e) Result obtained with final frequency of 0.5 MHz, 1.5 MHz, 2.5 MHz, and 3.5 MHz, respectively.

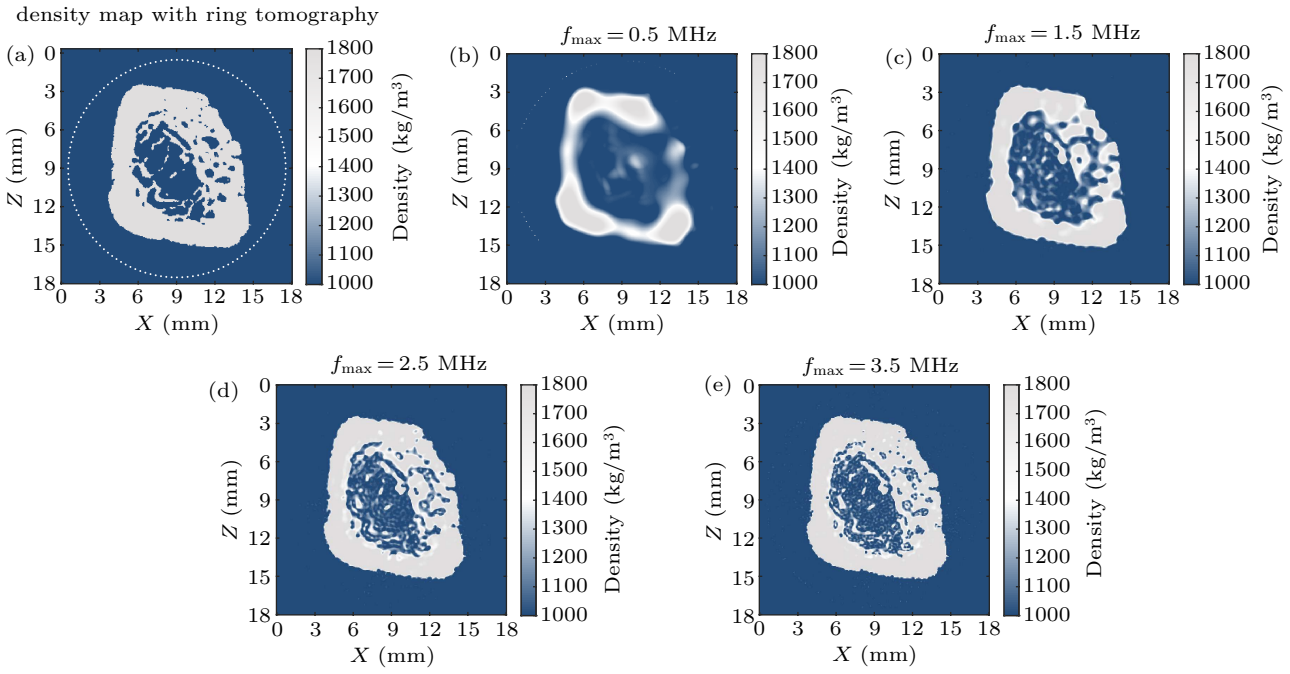


Fig. 6. FDFWI results of the single distal fibula for density model with different maximum frequencies. (a) True density model (derived from HR-pQCT image). (b)–(e) Result obtained with final frequency of 0.5 MHz, 1.5 MHz, 2.5 MHz, and 3.5 MHz, respectively.

relative error is 14.59%. For the mass density, RMSE in ROI is 150.96 kg/m^3, and the mean relative error in ROI is 10.78%. RMSE in the bone area is 226.74 kg/m^3, and the mean relative error is 12.60%. As presented in Fig. 7, the profile errors of the inversion results with the final frequency from 2.5 MHz to 3.5 MHz are similar, which further proves the limited improvement of the final images with the maximum frequency being higher than 2.5 MHz. Figure 8 shows that the inversion results are in good agreement with the most of structure and edge information extracted from the true model. Although the outer boundary of the density map is reconstructed well, the inaccurate recovery of microstructure in bone area can be observed. The inversion result of sound velocity is better than that of mass density.

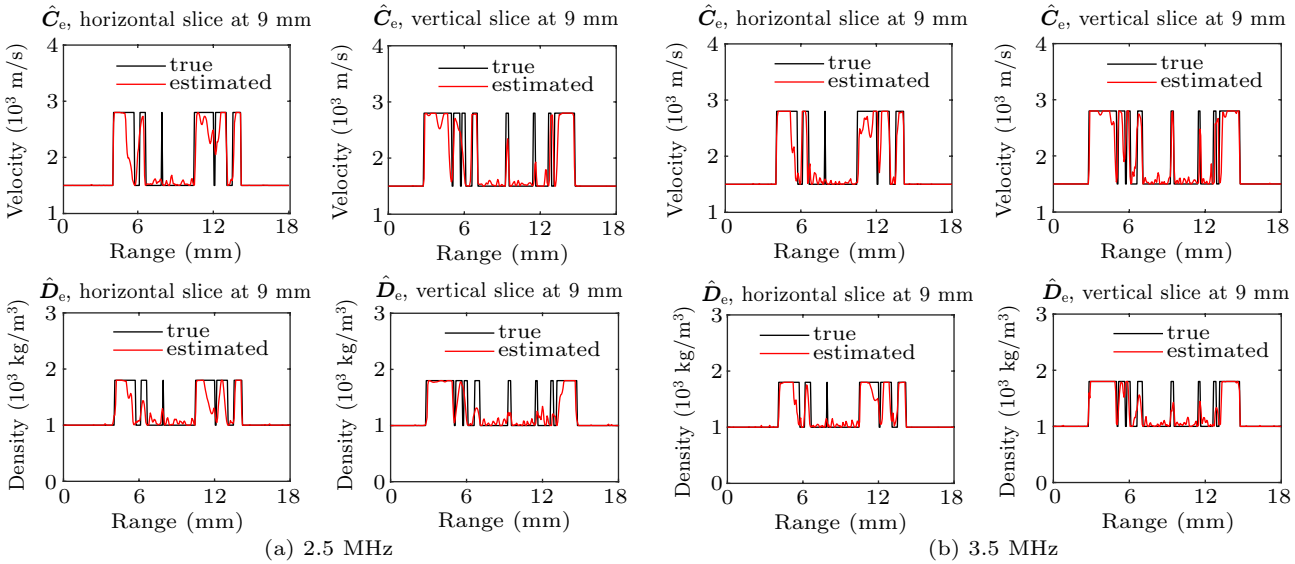


Fig. 7. Comparisons of profiles for velocity and density inversion results at $z = 9$ mm and $x = 9$ mm with those of the true velocity model in Fig. 5(a) and the true density model in Fig. 6(a) using maximum frequency of 2.5 MHz (a) and 3.5 MHz (b), respectively. The black and red lines are the profiles of the true model and the inversion result, respectively.

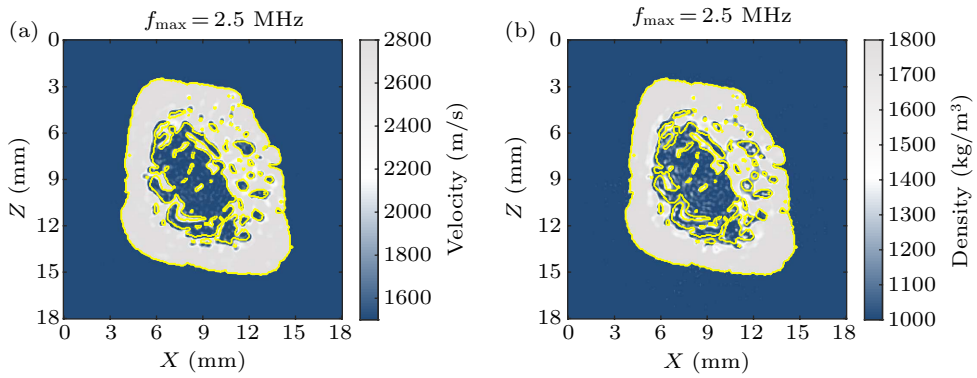


Fig. 8. Comparison structure and edge information between inversion results and true models with the maximum frequency 2.5 MHz. (a) Velocity inversion result. (b) Density inversion result. The yellow line represents structure and edge information extracted from the true model.

4.3. Distal tibia-fibula pair model

Figure 9 shows that the inversion result is in good agreement with the most of structure and edge information extracted from the true velocity model. The distribution of trabeculae in distal tibia is partly reconstructed and the estimated direction of trabeculae can match the true model. The result demonstrates that FDFWI has the potential to provide image of trabecular bones. At least partial trabecular distribution in wavelength scale can be provided. However, the reconstruction of density map in Fig. 10 is not as good as that of velocity map. For the sound velocity, RMSE in ROI is 217.39 m/s, and the mean relative error is 10.11%. RMSE in the bone area is 615.64 m/s, and the mean relative error is 21.99%. For the mass density, RMSE in ROI is 183.63 kg/m^3, and the mean relative error is 13.12%. RMSE in the bone area is 532.24 kg/m^3, and the mean relative error is 29.57%. The reconstruction errors of both sound velocity and mass density in the bone area are larger than those of the single distal fibula model. We consider that the multiple scattering and diffraction effects introduced by trabeculae in distal tibia bone increase the errors of inversion. This could also be the reason for the underestimation of bone parameters in the trabecular region. The phenomenon is depicted in Fig. 11. The inversion result shows that the velocity map can be recovered well even in such a complex sound field environment.

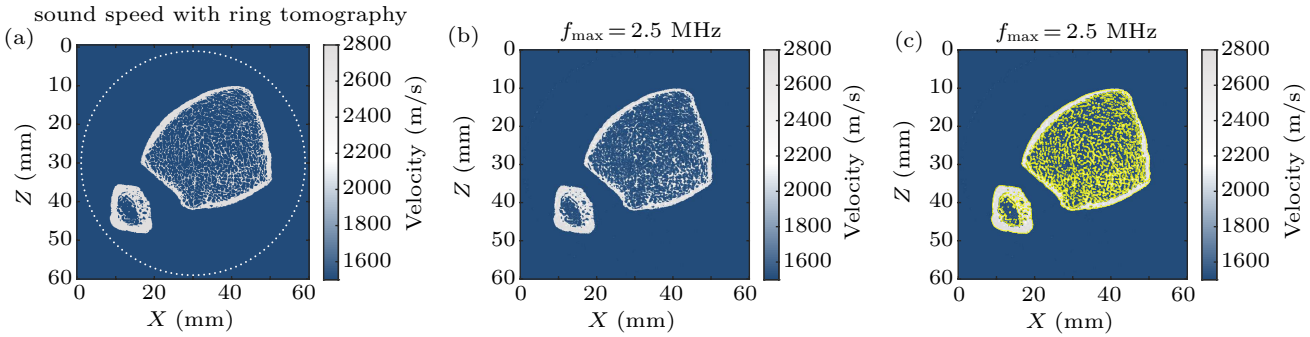


Fig. 9. FDFWI results of the distal tibia-fibula pair for velocity model with maximum frequency 2.5 MHz. (a) True velocity model (derived from HR-pQCT image). (b) Velocity inversion result. (c) Comparison structure and edge information between inversion result and true model. The yellow line represents structure and edge information extracted from the true model.

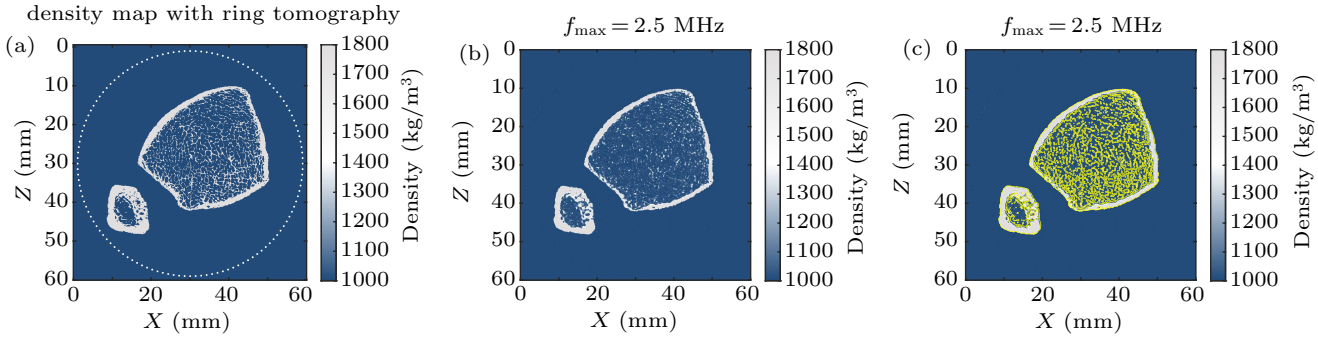


Fig. 10. FDFWI results of the distal tibia-fibula pair for density model with maximum frequency 2.5 MHz. (a) True density model (derived from HR-pQCT image). (b) Density inversion result. (c) Comparison structure and edge information between inversion result and true model. The yellow line represents structure and edge information extracted from the true model.

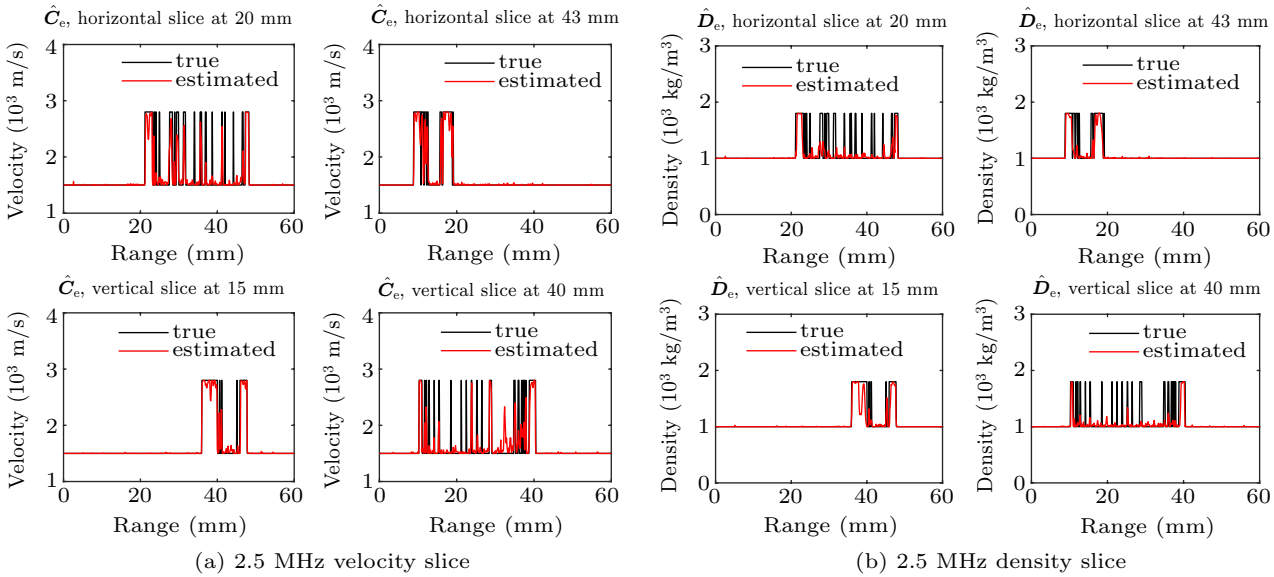


Fig. 11. Comparisons of profiles for inversion velocity and density results with those of the true velocity model in Fig. 9(a) and density model in Fig. 10(a). (a) Velocity profiles, (b) density profiles at $z = 20$ mm, $z = 43$ mm, $x = 15$ mm, and $x = 40$ mm. The black and red lines are the profiles of the true model and the inversion result, respectively.

4.4. Robustness against noise

As shown in Fig. 12, the sound velocity maps can still be recovered well despite the presence of noise. Even in the case of low SNR (i.e., 0 dB), the geometry of bone and some relatively larger microstructure can be reconstructed, but the trabecular direction is difficult to be identified. The results with noises display the robustness of the method to random Gaussian noise in velocity inversion. Figure 13 indicates that the noisy effect on density inversion is greater than that on velocity inversion, especially for the case with low SNR.

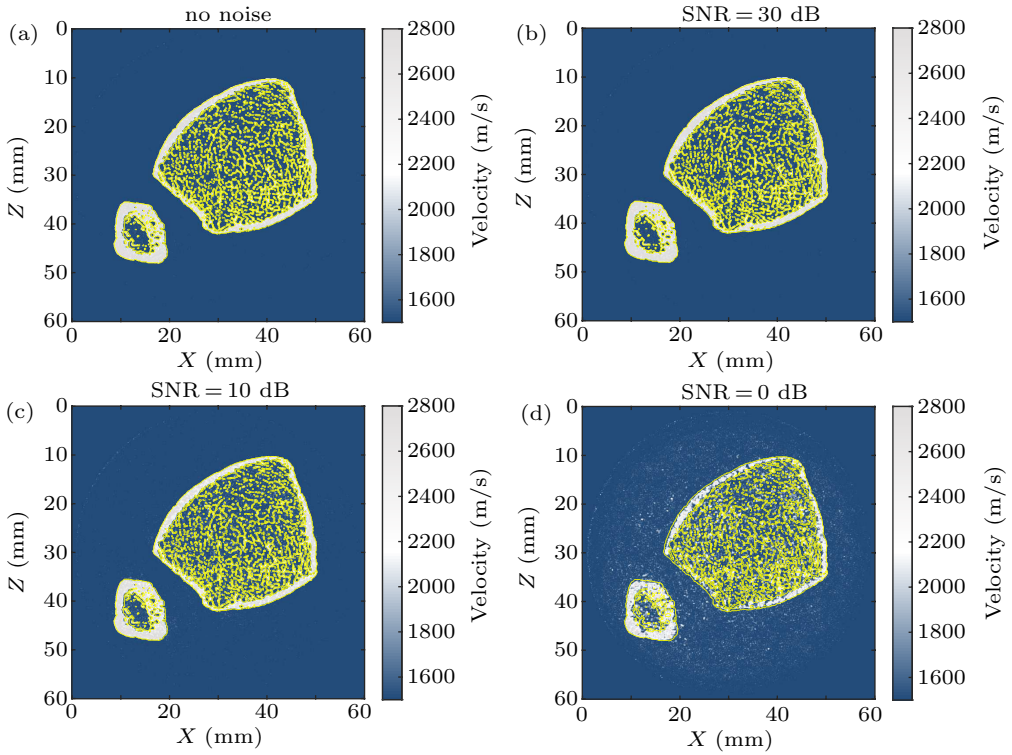


Fig. 12. Comparison structure and edge information between velocity inversion result and true model of the distal tibia-fibula pair with different SNRs. (a) No noise. (b) SNR = 30 dB. (c) SNR = 10 dB. (d) SNR = 0 dB. The yellow line represents structure and edge information extracted from the true model.

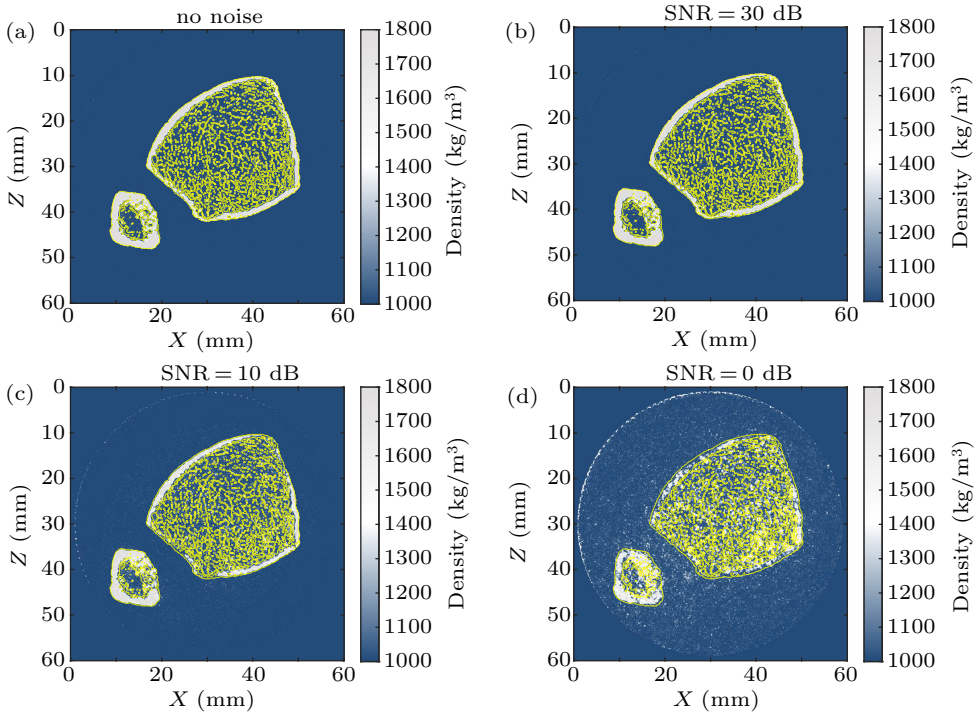


Fig. 13. Comparison structure and edge information between density inversion result and true model of the distal tibia-fibula pair with different SNRs. (a) No noise. (b) SNR = 30 dB. (c) SNR = 10 dB. (d) SNR = 0 dB. The yellow line represents structure and edge information extracted from the true model.

5. Discussion

In practice, ultrasonic waves in cortical bone travel as guided waves. The shear and longitudinal waves are coupled with each other. However, ultrasonic sources and receivers immersed in water do not generate or record shear waves. Meanwhile, elastic conversions are small close to normal incidence, which is the most important part of wavefield for FWI, both in transmission and reflection imaging.[39] The layout of ring array plus a nearly circular imaging plane (cross-section) can ensure relatively slight effects of elastic conversions.[39] From the perspective of accurate model matching, the closer the numerical forward model is to true bone and surrounding tissue,

the better the observed wavefield can be approximated. That is, a more complete account of the physics effects during inversion, including absorption, anisotropy, and elasticity in hard tissues, can theoretically provide a more quantitatively accurate model with physical properties, but at a cost of significant computation increasing to acquire the final model. A recent *in-vivo* FWI study[39] shows that a simple isotropic and acoustic model is adequate for providing well-resolved images for hard tissue (i.e., the skull) and soft tissue (i.e., the brain). The problem investigated in this paper is very similar to the issue that has been addressed successfully for brain imaging using FWI.[39] Consequently, the assumptions of the bone and surrounding soft tissue are acoustic and isotropic media. The simulated results show that the simply acoustic model can recover high-resolution image, including hard tissue (i.e., bone) and surrounding soft tissue (i.e., marrow, muscles, fat, and skin).

The spatial resolution of ultrasonic imaging is the minimum distance between two reflectors that can be distinguished. The axial resolution of the traditional pulse-echo method depends on pulse duration and is equal to half the spatial pulse length.[33] The lateral resolution is determined by the width of the ultrasonic beam.[33,34] The pulse-echo method can achieve good axial resolution at high frequencies using short pulse length. The lateral resolution can be improved capitalizing on synthetic aperture or focused ultrasound.[33,36] However, it is necessary to know the distribution of the velocity model or to estimate the sound velocity prior to bone imaging.[32] Otherwise, the imaged bone will be distorted and displaced due to the large velocity contrast. In addition, the advance estimation of sound velocity further increases the complexity of the algorithm.

For travel-time tomography, the spatial resolution is limited by the first Fresnel zone, whose diameter at the midpoint is $\sqrt{\lambda d}$, where λ is the wavelength, and d denotes the distance between the source and receiver pair.[39] The first Fresnel zone describes the region through which the ultrasonic wave can propagate from a source to a receiver, arriving within the same half cycle. Since the structure below the scale of the first Fresnel zone cannot change the observed delay time,[45] the travel-time tomography fails to give sub-Fresnel-zone structure of bone. When the emitter and receiver are far apart and the frequency is low, the diameter of the first Fresnel zone is relatively large. If the highest frequency is 2.5 MHz and the furthest distance between the emitter and receiver is 58 mm, the worst resolution is about 8 mm. In this case, the method cannot provide accurate thickness and the microstructure of the bone (i.e., the wavelength scaled pores and trabeculae).

FDFWI is not limited by the above factors. With this method, bone parametric images are obtained by optimizing the misfit between the numerically simulated and observed wavefield. The physical properties of bone, such as sound velocity and mass density, are estimated during the imaging process. As long as the wavefield can uniformly spread throughout the entire imaging area, FDFWI has the ability to distinguish the wavefield from the scattered points separated by more than half a wavelength in all directions.[58] Therefore, the spatial resolution in this work for bone tissue using the maximum frequency 2.5 MHz is about 0.6 mm. The resolution is small enough for imaging part of the bone microstructure. The requirement that the wavefield should cover the entire imaging region is due to a well-known formula (15) obtained in geophysics,[50] which links the observation angle to spatial frequency content of the reconstructed image,

$$|K| = \left| \frac{4\pi f}{C} \cos\left(\frac{\theta}{2}\right) \right|, \tag{15}$$

where K is the wavenumber, f is the frequency of the incident wave, and θ is the scattering angle, that is the angle between the source and the receiver. Bernard *et al.*[31] have demonstrated the importance of having a circular sampling of the wavefield in simulation study. It is not possible to reconstruct an accurate and high-resolution image using only transmitted or reflected data. That is the reason why we chose ring array transducers in this paper (i.e., scattering angles covering all directions).

If we compare the inversion results of sound velocity, as shown in Figs. 5 and 9, with binarized images derived from HR-pQCT (i.e., Figs. 5(a) and 9(a)), the FDFWI method can achieve high-resolution bone images with sub-millimeter which are close to those of HR-pQCT. However, the results are only reconstructed with numerical simulation, and the performances of *in-vitro* and *in-vivo* experiments need to be further verified in the future work.

The inverted velocity maps are very close to the true models, both in terms of geometry and velocity values, even for the distal tibia-fibula pair model with multiple reflections and diffractions between the two bones. At the same time, the multiple scattering and diffraction effects in trabecular region further pose a significant imaging challenge. Nevertheless, the velocity maps still can be reconstructed well, even with significant noise. In contrast, the density inversion accuracy is far less accurate. In Figs. 6 and 10, the artifacts of density maps in marrow cavities are more than those in the velocity maps. As shown in Fig. 10, the inversion results in part of the bone regions cannot accurately present the true thickness, and are underestimated to some extent. The poor recoveries of boundaries in the region, where the two bones are close to each other, can be observed. There are two main reasons to account for this phenomenon. One is that the multiple reflection and diffraction exist between the boundaries of the two bones. The other is that it is difficult for the receivers to directly obtain the reflected waves generated by the bone boundaries that

are adjacent to each other, for the reflected wavefields will be affected by the scattering wavefield induced by the trabeculae and cortical boundaries that are closer to the ring array transducer. The receivers get more information of transmitted wavefield rather than that of the reflected wavefield in this region. Nevertheless, the density inversion is sensitive to the reflection rather than the transmission wavefield.[50]

In geophysics, studies[50,75] have proved that simultaneous inversion of sound velocity and density is difficult. To obtain a relatively good estimation of density, either an extremely low initial frequency[76,77] or multiple times for inversion are required.[77] The sound velocity is sensitive to the data of various scattering angles, but the mass density is only interested in that of small scattering angle.[50] As a result of Eq. (15), only the wavefields sampled from reflections, which contain the high spatial frequency information, are available for mass density reconstruction. Since mass density is an important indicator in clinical bone disease diagnosis, investigating more effective methods for accurate density recovery is our future work.

As shown in the inversion results (i.e., Figs. 4–6), there exist some noises in the marrow cavity, which are similar to the noisy effects in study.[31] Two underlying causes could account for the effects. One is that the frequency-domain information for inversion using several frequency components is still limited rather than utilizing the complete frequencies. On the other hand, the noisy effect could represent an overestimate in some local area for parametric imaging. In a sense, it could be a local minimum that is very close to the global optimal solution. The noisy effect deserves further refinement in the future. Due to the reflection, diffraction, or scattering caused by the circular transducer, overestimates for some grids at the locations of the elements in the ring array can be observed in Fig. 6(b) for mass density inversion at initially low frequency components, but the phenomena remarkably improved with increasing frequency until the effects disappeared.

As depicted in Fig. 12, even though the SNR is 0 dB, the geometries of bones can be reconstructed in velocity inversion. By comparing Figs. 12(a) and 13(a) with Figs. 12(b) and 13(b), the image quality of the noisy case with the SNR of 30 dB is comparable to that of the noiseless one. In the case of 10 dB in Fig. 12(c), the velocity inversion results are still good, and part of trabeculae and pores can be restored. As discussed above, accurate density inversion has been proved to be difficult, whereas figures 13(c) and 13(d) verify that significant noises may further complicate this task. With the SNR of 0 dB, there is much more noise in the density inversion than that in the velocity inversion. Therefore, the results show that the FDFWI is robust to random Gaussian noise in velocity inversion. The robustness comes from that the FDFWI method only approximates the wavefield information in the measured signals by calculating the full wave equation rather than producing noise. Therefore, the presence of noise may slow down the convergence to the true model, but the physical principle of FDFWI leads to its strong immunity to noise and the ability to reconstruct bone images at a significant noise level. This further indicates that the proposed FDFWI method has the potential to adapt to the low SNR experimental environment.

The omni-directional point source in the simulation was different from the experimental transducer element in the size, beam width, and direction, which might lead to the differences of the excited waveforms between the simulation and experiment. The differences between the simulated and experimental sources could further influence the inversion processes. To best match the observed wavefields, the setting of the element size in simulation should be consistent with that in the experiment, and the simulated sources could be defined as the experimental measured excited waveforms of the elements in the ring array. Another kind of the source calibration and estimation techniques can refer to the reference.[61]

In practical application, it is difficult to obtain a transducer with bandwidth from 100 kHz to 2.5 MHz, especially covering the low frequency with sufficient SNR. There are several solutions to address this issue. One is that the results derived from travel-time[54] or adaptive waveform inversion[39] can be used as the initial model for inversion, which is closer to the real model. In this case, the closer the model is to the real model, the higher the initial frequency can be used. The second way is to utilize a wide-band transducer, which now has been available to customize a wide-band transducer with multiple center frequency coupled. The third method is that the inversion can be performed using multiple transducers with the same size and different dominant frequencies. In this study, two sources with different center frequencies were used to cover the bandwidth we need.

Ultrasonic attenuation in bone may be resulted from a number of factors, including boundary reflections, ultrasonic diffraction, scattering, and reverberation effects introduced by porosity and trabeculae, anelastic absorption in the bone and surrounding tissues, and elastic mode-conversions. In this study, the anelastic absorption and elastic mode-conversions were not considered. This model assumption underestimated the attenuation effect in bone tissues, which might degrade the imaging quality of bone structure for experimental cases. Meanwhile, it should be noted that the frequency-dependent attenuation coefficients in different individuals are quite different. It is necessary to dynamically adjust the maximum frequency according to the actual situation. If the attenuation coefficients in the bone and soft tissues are significantly larger than the mean attenuations,[78–80] the maximum frequency should be decreased to a lower value to guarantee sufficient SNR. The decrease of the maximum frequency may influence

the resolution and the image quality. As aforementioned, if the efficient computation is available, a more complete consideration of the physics effects for forward modeling is worth further investigation in the future.

To facilitate modeling, the theoretically parametric models (i.e., sound velocity and mass density) were constructed using binary values in this study. The bone and surrounding soft tissue were assumed to be homogeneous media. In clinical applications, both soft tissue and bone may exhibit heterogeneity, which represent as multivalued parameter models. The influence of surrounding soft tissues on QUS based assessment of bone properties is a long-standing problem.[13,81] It is not easy for some quantitative approaches, such as guided wave, backscatter to accurately consider the influence of heterogeneity in soft tissue. However, theoretically, this is not a problem for FDFWI, which iteratively updates the physical parameters in the grids of the simulated model during the optimization process until the best match reaches. The FDFWI method does not need to explicitly differentiate the bone and soft tissue, nor does it require the additional processing of heterogeneity within the same tissue. The properties of bone and surrounding soft tissues can be inverted jointly. The ultrasound propagation through different tissues or different regions of the same tissue will carry specific features, which will be utilized to precisely characterize the model parameters in inversion. Both high-contrast and low-contrast areas can be accurately reconstructed.

Unlike TDFWI based on the iterative matching of complete time series between simulated and observed data, the FDFWI only depends on the frequency-domain wavefields at limited discrete frequencies. This data reduction improves calculation efficiency. In Ref. [31], a model with the dimension of 667×667 took about an hour on a powerful desktop PC (specific conditions were not provided) to obtain the final inversion result, utilizing only 5 frequency bands (12 min/frequency band). For a parametric model with the dimension of 601×601 in this study, it took approximately one and a half hours to get the final results using 49 discrete frequency components (1.84 min/frequency component, core Intel Core CPU i9-9900K@5.10 GHz, DDR 64 GB, without graphic processing unit GPU). Compared to TDFWI, the computation efficiency of FDFWI is obviously improved. For clinical application, FDFWI for three-dimensional (3-D) imaging has the potential to be performed less than 10 min using parallel computation,[39] distributing one source to one GPU in virtue of several GPU-pool based servers to further speed up the inversion process. The FDFWI might meet the real-time requirement for medical imaging with the improvement in computation in the future.

6. Conclusion and perspectives

In this study, USCT based on FDFWI has been applied in 2-D bone quantitative imaging. The simulated results show that the method can achieve high-resolution parametric imaging. Not only macroscopic morphology but also microstructure (i.e., wavelength scaled pores and trabeculae) can be reconstructed clearly and accurately in the velocity map with sub-millimeter resolution. The performance of the velocity map is close to that of HR-pQCT in numerical simulation. Although restoring high-precision density maps using FDFWI has also been proved to be a difficult task, the method can still restore the macroscopic morphology and part of the microstructure in density map. The method has also been demonstrated being robust against Gaussian noise in velocity inversion. Therefore, the proposed FDFWI method has the potential for the diagnosis and daily monitoring of bone disease. It can also be applied to image multilayer and irregular objects with high-contrast impedance in industrial nondestructive testing. Future work will focus on experimental data instead of synthetic data to further demonstrate the effectiveness of the method.

References

[1] Seeman E 2008 *J. Bone Miner Metab* **26** 1
[2] Schneider J, Ramiandrisoa D, Armbrecht G, Ritter Z, Felsenberg D, Raum K and Minonzio J G 2019 *Ultrasound Med. & Biol.* **45** 1234
[3] Minonzio J G, Bochud N, Vallet Q, Bala Y, Ramiandrisoa D, Follet H, Mitton D and Laugier P 2018 *Bone* **116** 111
[4] Compston J E, McClung M R and Leslie W D 2019 *Lancet* **393** 364
[5] Kanis J A 2002 *Lancet* **359** 1929
[6] Matsukawa M 2019 *Jpn. J. Appl. Phys.* **58** SG0802
[7] Liu C, Li B, Diwu Q, Li Y, Zhang R, Ta D and Wang W 2018 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **65** 2311
[8] Liu C, Li B, Li Y, Mao W, Chen C, Zhang R and Ta D 2020 *Ultrasound Med. & Biol.* **46** 305
[9] Liu C, Dong R, Li B, Li Y, Xu F, Ta D and Wang W 2019 *Chin. Phys. B* **28** 024302
[10] Donnelly E 2011 *Clin Orthop Relat. Res.* **469** 2128
[11] Lasaygues P, Guillermin R and Lefebvre J P 2011 *Bone Quantitative Ultrasound* 1st edn. (Dordrecht: Springer) pp. 441–459
[12] Minonzio J G, Bochud N, Vallet Q, Ramiandrisoa D, Etcheto A, Briot K, Kolta S, Roux C and Laugier P 2019 *J. Biomed Mater Res.* **34** 1585
[13] Bochud N, Vallet Q, Minonzio J G and Laugier P 2017 *Sci. Rep.* **7** 1
[14] Okumura S, Nguyen V H, Taki H, Haïat G, Naili S and Sato T 2018 *Appl. Sci.* **8** 652
[15] Moilanen P, Kilappa V, Nicholson P H F, Timonen J and Cheng S 2004 *Ultrasound Med. & Biol.* **30** 1517
[16] Padilla F, Jenson F, Bousson V, Peyrin F and Laugier P 2008 *Bone* **42** 1193
[17] Hoffmeister B K, McPherson J A, Smathers M R, Spinolo P L and Sellers M E 2015 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **62** 2115
[18] Wear K A 2020 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **67** 454
[19] Denis M, Wan L, Fatemi M and Alizad A 2018 *Ultrasound Med. & Biol.* **44** 714
[20] Mohanty K, Yousefian O, Karbalaeisadegh Y, Ulrich M, Grimal Q and Muller M 2019 *Comput. Biol. Med.* **114** 1
[21] Lahivaara T, Karkkainen L, Huttunen J M J and Hesthaven J S 2018 *J. Acoust Soc. Am.* **143** 1148
[22] Li Y, Xu K, Li Y, Hu B, Zhang J, Le L H and Ta D 2019 *IEEE International Ultrasonics Symposium (IUS)* Oct, Glasgow
[23] Foiret J, Minonzio J G, Chappard C, Talmant M and Laugier P 2014 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **61** 1478

[24] Bochud N, Vallet Q, Bala Y, Follet H, Minonzio J G and Laugier P 2016 *Phys. Med. Biol.* **61** 6953

[25] Schneider J, Iori G, Ramiandrisoa D, Hammami M, Grasel M, Chappard C, Barkmann R, Laugier P, Grimal Q, Minonzio J G and Raum K 2019 *Arch Osteoporos* **14** 21

[26] Okumura S, Nguyen V H, Taki H, Haïat G, Naili S and Sato T 2017 *Jpn. J. Appl. Phys.* **56** 07JF06

[27] Jensen J A, Nikolov S I, Gammelmark K L and Pedersen M H 2006 *Ultrason.* **44** Suppl 1 e5

[28] Garcia D, Le T L, Muth S, Montagnon E, Poree J and Cloutier G 2013 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **60** 1853

[29] Couture O, Hingot V, Heiles B, Muleki-Seya P and Tanter M 2018 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **65** 1304

[30] Errico C, Pierre J, Pezet S, Desailly Y, Lenkei Z, Couture O and Tanter M 2015 *Nature* **527** 499

[31] Bernard S, Monteiller V, Komatitsch D and Lasaygues P 2017 *Phys. Med. Biol.* **62** 7011

[32] Renaud G, Kruizinga P, Cassereau D and Laugier P 2018 *Phys. Med. Biol.* **63** 125010

[33] Jiang C, Li Y, Li B, Liu C, Xu F, Xu K and Ta D 2019 *IEEE Access* **7** 163013

[34] Li Y, Jiang C, Li Y, Xu F, Xu K, Ta D and H L L 2019 *Acta Phys. Sin.* **68** 184302 (in Chinese)

[35] Olofsson T 2010 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **57** 2522

[36] Wu H, Chen J, Yang K and Hu X 2016 *Meas. Sci. Technol.* **27** 045401

[37] Qin K, Yang C and Sun F 2014 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **61** 133

[38] Wu S and Yang K 2014 *IEEE International Ultrasonics Symposium (IUS)* Sep, Chicago

[39] Guasch L, Calderon A O, Tang M X, Nachev P and Warner M 2020 *npj Digit. Med.* **3** 1

[40] Lasaygues P, Ouedraogo E, Lefebvre J P, Gindre M, Talmant M and Laugier P 2005 *Phys. Med. Biol.* **50** 2633

[41] Li Y, Li B, Li Y, Liu C, Xu F, Zhang R, Ta D and Wang W 2019 *Ultrason Imaging* **41** 271

[42] Liu Q and Gu Y J 2012 *Tectonophysics* **566–567** 31

[43] Li C, Duric N, Littrup P and Huang L 2009 *Ultrasound Med. & Biol.* **35** 1615

[44] Qu X, Azuma T, Imoto H, Raufy R, Lin H, Nakamura H, Tamano S, Takagi S, Umemura S I, Sakuma I and Matsumoto Y 2015 *Jpn. J. Appl. Phys.* **54** 07FH10

[45] Rao J, Ratassepp M and Fan Z 2016 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **63** 737

[46] Rao J, Ratassepp M and Fan Z 2017 *J. Sound Vib.* **400** 317

[47] Guillermin R, Lasaygues P, Rabau G and Lefebvre J P 2013 *J. Acoust Soc. Am.* **134** 1001

[48] Zheng R, Le L H, Sacchi M D and Lou E 2015 *Ultrasound Med. Biol.* **41** 2955

[49] Perez-Liva M, Herraiz J L, Udias J M, Miller E, Cox B T and Treeby B E 2017 *J. Acoust Soc. Am.* **141** 1595

[50] Virieux J and Operto S 2009 *Geophysics* **74** WCC1

[51] Operto S, Gholami Y, Prieux V, Ribodetti A, Brossier R, Metivier L and Virieux J 2013 *Lead. Edge* **32** 1040

[52] Li C, Huang L, Duric N, Zhang H and Rowe C 2009 *Ultrasonics* **49** 61

[53] Hooi F M and Carson P L 2014 *Med. Phys.* **41** 082902

[54] Wang K, Matthews T, Anis F, Li C, Duric N and Anastasio M A 2015 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **62** 475

[55] Falardeau T and Belanger P 2018 *J. Acoust Soc. Am.* **144** 2937

[56] Lasaygues P, Rouyer J, Mensah S, Franceschini E, Rabau G, Guillermin R, Bernard S, Monteiller V and Komatitsch D 2017 *International Workshop on Medical Ultrasound Tomography* Nov, Speyer pp. 77–88

[57] Lu C, Lin J, Chew W and Otto G 1996 *Ultrason Imaging* **8** 140

[58] Pratt R G 2017 *International Workshop on Medical Ultrasound Tomography* Nov, Speyer pp. 65–76

[59] Pratt R G 1999 *Geophysics* **64** 888

[60] Pratt R G and Shipp R M 1999 *Geophysics* **64** 1942

[61] Sandhu G Y, Li C, Roy O, Schmidt S and Duric N 2015 *Phys. Med. Biol.* **60** 5381

[62] Pratt R G and Worthington M H 1990 *Geophys. Prospect.* **38** 287

[63] Demmel J W, Eisenstat S C, Gilbert J R, Li X S and Liu J W H 1999 *SIAM J. Matrix Anal. Appl.* **20** 720

[64] Hormati A, Jovanović I, Roy O and Vetterli M 2010 *Med. Imaging 2010: Ultrasonic Imaging, Tomography and Therapy* Mar, San Diego

[65] Plessix R E 2006 *Geophys. J. Int.* **167** 495

[66] Tromp J, Komatitsch D and Liu Q 2008 *Commun. Comput. Phys.* **3** 1

[67] Wang Q, Zhang J and Huang Z 2015 *Prog. Geophys. (Chin.)* **30** 2797

[68] Byrd R H, Lu P, Nocedal J and Zhu C 1995 *SIAM J. Sci. Comput.* **16** 1190

[69] Kalita M, Kazei V, Choi Y and Alkhalifah T 2019 *Geophysics* **84** R569

[70] Lin Y and Huang L 2015 *Geophys. J. Int.* **203** 2125

[71] Lin Y and Huang L 2014 *Geophys. J. Int.* **200** 489

[72] Kazei V, Kalita M and Alkhalifah T A 2017 *79th EAGE Conference and Exhibition* 2017 Jun, Paris, p. 1

[73] Bernard S, Schneider J, Varga P, Laugier P, Raum K and Grimal Q 2016 *Biomech Model Mechanobiol* **15** 97

[74] Xu K and McMechan G A 2014 *Geophysics* **79** R41

[75] Tarantola A 1986 *Geophysics* **51** 1893

[76] Choi Y, Min D J and Shin C 2008 *Geophys. Prospect.* **56** 863

[77] Jeong W, Lee H Y and Min D J 2012 *Geophys. J. Int.* **188** 1221

[78] Haïat G 2011 *Bone Quantitative Ultrasound* 1st edn. (Dordrecht: Springer) pp. 331–360

[79] Wear K A 2001 *IEEE Trans. Ultrason Ferroelectr. Freq. Control* **48** 602

[80] Pakula M, Padilla F and Laugier P 2009 *J. Acoust Soc. Am.* **126** 3301

[81] Moilanen P, Salmi A, Kilappa V, Zhao Z, Timonen J and Hæggström E 2017 *J. Appl. Phys.* **122** 144901